

# FUNCTIONAL ANNOTATION OF EXPRESSED SEQUENCE TAGS (ESTs) IN *ARABIDOPSIS THALIANA*

**Manoj Kumar Sharma, Astha Duhan\* and Rajeshwari Sharma**

\*Department of Bioinformatics, J.V. College, Baraut, Baghpat (U.P.)

\*\*Department of Bioinformatics, Mata Gujri College, Fatehgarh Sahib (Punjab).

**Abstract:** Expressed Sequence Tags are the short single pass reads obtained from the ends of a cDNA clone. They are currently the most widely sequenced moiety in case of nucleotide sequences. ESTs provide a cost effective way to gene indexing, gene discovery and the characterization of alternatively spliced forms. The information of what gene EST belongs to or is it coming from a new gene can be easily done by clustering the EST sequences and assembling them into contigs, indexing them such that each contig contains the information for only one gene. The present study was aimed at the same principle, following which, 3'-ESTs from *Arabidopsis thaliana* were clustered and a total of 4018 contigs were generated. Of the total 31,518 ESTs under study, 17,853 remained unassembled and were marked as unassembled or Singleton ESTs. Among the 4018 contigs, 3767 contigs gave a match with the nonredundant database with a success rate of 93.75%.

**Key words:** EST, CODONCODE ALIGNER, NCBI, BLAST.

## REFERENCES

- Adams, M.D.; Kelley, J.M.; Gocayne, J.D.; Dubnick, M.; Polyneropoulos, M.H.; Xiao H.; Merrill, C.R.; Wu, A.; Olde, B.; Moreno, R.F.; Kerlavage, A.R.; McConbie, W.R. and Venter, J.C. (1991). Complementary DNA Sequencing: Expressed Sequence Tags and Human Genome Project, *Science*, **252**: 1651-1656.
- Boguski, M.S. *et al.* (1993). dbEST - database for 'expressed sequence tags', *Nat. Genet.*, **4**:332-333.
- Rudd, S. *et al.* (2003). Sputnik: a database platform for comparative plant genomics. *Nucleic Acids Res.* **31**: 128-132.
- Yu, J.; Whitelaw, C.A.; Nierman, W.C.; Bhatnagar, D. and Cleveland, T.E. (2004). *Aspergillus flavus* expressed sequence tags for identification of genes with putative roles in aflatoxin contamination of crops, *FEMS Microbial Lett* **237**: 333-340.